### Interneurons: their cognitive roles

### - A perspective from dynamical systems view

### Hiroshi Fujii

Department of Information & Communication Sciences

Kyoto Sangyo University

Kyoto 603-8555, Japan

fujii@bacco.kyoto-su.ac.jp

Abstract - Interneurons may function as coordinator of cortical dynamics, as "super networks" driven by top down signals. Dynamical systems-theoretic view on cortical dynamics and its hypothetical cognitive roles are discussed.

In particular, we discuss three related issues concerning the cognitive role of inhibitory interneurons in the neocortex.

- Theoretical possibility of emergence of transient synchrony in chaotic itinerant dynamics in gap junctioncoupled interneuron networks.
- 2. Dynamical systems-theoretic, and non-classical memory view based on the concept of Milnor attractor.
- A new hypothesis on functional role of activity-dependent depolarizing action of GABA<sub>Λ</sub> in inhibitory interneurons.

Index Terms – interneuron, super network, transient synchrony, gap junction, chaotic itinerant dynamics, Milnor attractor, depolarizing action of  $GABA_A$ 

#### I. Introduction

Interneurons in the neocortex and the hippocampal formation have once been regarded as something which plays auxiliary roles in the organization of brain activity and consequently in cognitive functions. Interneurons, mostly inhibitory ones, are believed to be designed to prevent instability of the brain system by inhibition. It is sometimes believed also that interneurons form local or micro circuitries in the neocortex, and are fed locally by nearby "principal" (pyramidal) neurons.

However, recent experimental data suggest that such a picture on interneurons, at least for some subsystems, should undergo a fundamental change. Interneurons may play essential roles in the organization of brain activity through interplay with pyramidal systems.

Interneurons in the neocortex and hippocampus exhibit a large diversity in morphology, firing properties and biochemical patterns of coexpression [7], [8], [19] and may form discrete networks, although their classification is largely controversial.

### Ichiro Tsuda

Department of Mathematics, Graduate School of Science
Hokkaido University
Sapporo 060-0810, Japan
tsuda@math.sci.hokudai.ac.jp

The possibility exists that certain sub-networks of interneurons might be driven by top down signals as *context*, *attention* or e.g., *consciousness*, and which in turn control and cooperatively entrain pyramidal networks. In fact, it is reported that the action of single perisomatic inhibitory cells can synchronize 1,000-2,000 principal cells [7]. Buzsaki [4] suggests that networks of inhibitory interneurons might form 'super networks', and which impose a coordinated oscillatory "context" for the "content" carried by principal cells.

Together with increasing evidence about the possible role of interneurons in the genesis of rhythms and transient synchrony, study of activity and functions of "super networks" by interneurons, and their supra and subthreshold influence on pyramidal systems, would be an important step to establish a unified view on the cortical dynamics and its functional significance.

We shall discuss three related issues under such a perspective concerning the cognitive role of inhibitory interneurons from dynamical systems standpoints.

- 1. Theoretical possibility of emergence of transient synchrony in chaotic itinerant dynamics in gap junction-coupled interneuron networks.
- Dynamical systems-theoretic, and non-classical memory view based on the concept of Milnor attractor.
- A new hypothesis on functional role of activitydependent depolarizing action of GABA<sub>A</sub> in inhibitory interneurons.

# II. GAP JUNCTIONS MAY ORGANIZE TRANSIENT SYNCHRONY THROUGH CHAOTIC ITINERANCY

The first issue we would like to discuss is the dynamical nature of gap junction-coupled systems of neurons.

Gap junctions (GJs), or electrical synapses, have been reported to exist among inhibitory interneurons in the

neocortex. GJs among neocortical interneurons are dendrodendritic. They are not only ubiquitous, but also *massive* in their connectivity as compared with chemical synapses ([12], [13], [2]). There are several distinct sub-networks coupled by gap junctions, and some of them possess chemical synaptic connections as well. However, the question about the GJ-connectivity among sub-classes of GJ-coupled FS (fast spiking) neurons exhibiting, e.g., distinct calcium-binding proteins or neuropeptides expressions is still not clear. They might be interconnected by GJs, or they might form discrete sub-networks driven by distinct origins of projections.

Although it is generally believed that (GJ-coupled) "interneurons generate a variety of synchronous inhibitory rhythms in the neocortex..." (J. R. Gibson *et al.*, 1999 [13]), little is known how those interneurons collectively behave when coupled by GJs in a massive way.

As is opposed to a general intuition that GJs organize synchrony among neurons, interneurons may, when coupled by GJs, exhibit a spatio-temporal chaos in their membrane potentials and spike trains as well, and *synchrony appears transiently within the itinerant dynamics*. Note that this phenomenon depends on the nonlinearity defined by ion channel properties of the concerned neuron. More precisely, if a cell possesses a property classified as *class* I\* as defined below, such itinerant states appear in a robust manner even under spatially uniform and temporally constant external environments, while they show only regular repetitive firings when they are isolated.

The model we consider here is based primarily on a *reduced form* of two-variables, where the first variable, say V, may represent the membrane potential, and the second one, R, the "recovery" variable representing an activation state of, e.g., some potassium channels in a generalized sense [24], [26]. With the injected current strength being denoted by i, the single cell equation may be written as:

$$\begin{cases} \tau_{V} \frac{dV}{dt} = f(V, R) + i \\ \tau_{R} \frac{dR}{dt} = g(V, R) \end{cases}$$
 (1)

According to the note made by Hodgkin in 1948 [17] and the fact that most of cortical neurons are of class I in the sense of Hodgkin (Cauli *et al.*, [5]), the two nullclines  $N_V = \{(V,R) | f(V,R)+i=0\}$  and  $N_R = \{(V,R) | g(V,R)=0\}$  are assumed to be of *class* I. Namely,  $N_R$  takes essentially *quadratic-like* form as compared with linear-like form for class II neurons. See, Fig.1. An instance of such an essentially quadratic nullcline is due to the introduction of slowly inactivating potassium currents. See, Rose and Hindmarsh [26]. The essence of this classification may become clear if one considers the mathematical mechanism of action potential generations: *homoclinic (saddle-node) bifurcations* for class I neurons, while *subcritical Hopf bifurcation for* class II neurons. A neuron of class II exhibits

a sudden and high frequency spike firing when the externally injected current i exceeds its threshold, which is a behaviour typically observed in the Hodgkin-Huxley neurons. On the other hand, a class I neuron with homoclinic bifurcations shows a gradually increasing firing rate as i exceeds its threshold.

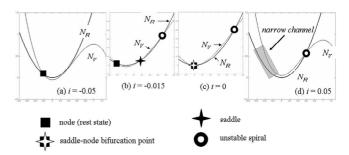


Fig.1. Class I\* neuron model in two-variable reduced form. The pictures (a) to (d) show the phase diagram of the  $\mu$ -model (Eqs.(1) and (2)) as the injected current i traverses the critical value i = 0. The bold and thin curves indicate respectively the nullclines  $N_R$  and  $N_V$ , respectively. Note the essentially quadratic form of the nullcline  $N_R$ which characterizes the class I neurons. Note also that class II neurons have essentially linear nullclines  $N_R$ . (Not shown here.) As the current i increases from the rest state (a), the two nullclines become to have three intersections, i.e., a stable node (which corresponds to a rest state), a saddle and an unstable spiral (see, (b)). Then, the stable node and the saddle collide at a critical current, where a homoclinic orbit appears from the saddle-node (c), which is the seed of action potentials (spikes) with the period  $= \infty$  . What characterizes the class I\* neurons is that as i traverses the saddle-node point, there appears a narrow channel between the two nullclines (d). (See, Fujii-Tsuda [10], 2004 for details.)

Let us introduce the concept of a *class* I\* neuron following the definition in [10].

Let I\* denote a subclass of class I neurons, in which there is an open interval  $\mathcal{G}=(i^*, i^{**})$  of injected currents such that for all  $i \in \mathcal{I}$ , it holds that

1. there is a family of closed orbits  $\mathcal{O}(i)$  with periods  $\Pi(i)$  such that

$$\lim_{\substack{i \to i^*}} \Pi(i) = +\infty$$

- 2. a *narrow channel* appears between the two nullclines for  $i > i^*$ , and
- 3. an unstable *spiral* fixed point exists in O(i).

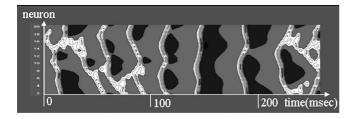
The above definition appears subtle, but this is merely a mathematical abstraction of the nonlinearity in, e.g., the Hindmarsh-Rose model [16], and our  $\mu$ -model illustrated in Fig.1:

$$f(V,R) = -R - \mu V^{2} (V - \frac{3}{2}), and$$
  

$$g(V,R) = -R + \mu V^{2}.$$
 (2)

We claim that if the neurons are of class I\*, the arguments which follow on the emergence of spatio-temporal chaos hold as well. Note also that a simple model of class I\* neuron with piecewise linear nullclines could serve as a model system for theoretical studies [10].

Fig.2 shows the contour map of membrane potentials of 20 neurons of class I\* (piecewise linear) neurons coupled with the two neighbours by GJs. Although each neuron receives a constant injected current without any external fluctuations, and no structural irregularity exists in the model, there appears an extensive fluctuation of membrane potentials with *transiently* synchronized states.



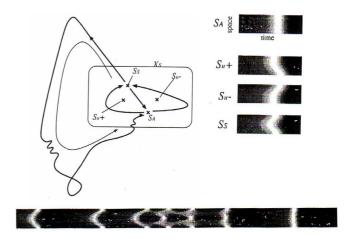
**Fig.2.** Contour map of membrane potentials (i.e., V in Eq. (1)) of 20 class  $I^*$  neurons of piecewise linear model coupled with the two neighbours by gap junctions with the coupling strength  $g_{GJ}=0.2$ . The vertical and horizontal directions indicate respectively the neuron positions, and the time (0->250 msec). Each neuron receives a constant injected current without any external fluctuations. No structural irregularity exists in this model. (From Fujii-Tsuda [9], 2004)

## A Milnor Attractor Controls the Itinerant Chaotic Dynamics

What is the mathematical nature of the chaotic dynamics revealed in Fig.2? To our surprise, our numerical study suggests, though mathematical verification remains to be done, that the concerned dynamics contains no (classical) attractors, but a unique (quasi-) attractor in the sense of Milnor [21]. By definition, a Milnor attractor possesses a positive measure of attracting region, but simultaneously may possess repelling orbits from the attractor. A Milnor attractor may thus provide a mechanism of allowing both transition and re-injection to and from a state.

The picture shown in Fig.3 illustrates the Poincaré map obtained from a system of GJ-coupled  $\mu$ -neurons exhibiting transient synchrony [32]. The Milnor attractor has a complicated internal structure connecting a few saddles which express particular symmetry patterns  $S_S$ ,  $S_A$ , .... The transient synchrony is a state near these saddles, and sometimes the orbit escape out of the Milnor "centre" exhibiting chaotic behaviour.

We emphasize again that GJs may organize synchrony, but that synchrony could be the one which appears transiently within the itinerant dynamics.



**Fig.3.** A fundamental structure of dynamical states for the transition between synchronized and desynchronized states.

Various dynamical states on the Poincaré section  $\Sigma_i V_i/N = 0.2$  are schematically drawn. This picture is a result of numerical computations in a system of GJ-coupled  $\mu$ -neurons (N=30) with i=0.00385, exhibiting transitions among various synchronous states  $S_S$ ,  $S_A$ ,..., where the corresponding spatio-temporal patterns are shown in the top right inset. The spatio-temporal pattern in the lowest inset can be represented by a Milnor attractor for the present chaotic itinerancy. (Modified from Tsuda et al. [32], 2004.)

### III. MEMORIES: BEYOND THE CONCEPT OF CLASSICAL ATTRACTOR VIEW

### Dynamical Systems-theoretic View of Memory Based on the Concept of Milnor Attractor

The study introduced above may provide us a theoretical motivation to a contemporary view on cortical dynamics. It may also provide an interpretation of *experimentally observed transient synchrony* in LFPs (local field potentials) and action potentials in the neocortex [14].

In fact, the study introduced in the above demonstrated the reality of dynamical systems consisted of coupled neurons where no classical attractors exist, while only attractors in Milnor's sense reside. When an orbit approaches a Milnor attractor (MA), the dynamics exhibits a synchronous state transiently, and stays there for a while.

Such a view on cortical dynamics may lead us to a new theoretical paradigm for the description of neocortical dynamics on which higher cognitive functions such as memory, perception, attention, ... are carried. It may provide us a unified view on basic questions such as how representation of information, as well as the dynamical mechanism of *binding*, is performed in the brain.

In fact, classical attractor view on the memory in the brain, such as the Hopfield theory and others, has a number of difficulties in, e.g., the description of episodic memories, together with the interpretation of physiological data on sub-threshold neocortical dynamics as *rhythm and synchrony genesis* which might be carried by inhibitory interneurons. (See, however e.g., Wang [34].)

The classical attractor view on the memory is based on the assumption that memories become existent through sculpting by learning in the brain. On the other hand, the contemporary view cannot assure the existence of memories as fixed objects, but assures memories only via dynamical recalling process.

We have presented a working hypothesis on how memories reside in the brain and related hypotheses. It may be summarized as:

- 1. Memories could be regarded as Milnor attractors (MAs) rather than classical *geometric* attractors.
- 2. Approaching of the orbit to a MA may be interpreted in the physiological term as the dynamical formation of a cell group which exhibits *synchrony* transiently.
- 3. The global dynamics in a *default* state may exhibit a chaotic itinerancy among MAs as memories ([11], [32], [33]]).
- It is not true to state that the recall is possible based on memories. On the contrary, memories can emerge through recalls.

We note that such viewpoints have long been proposed by I. Tsuda [30], [31] with the concept of "chaotic itinerancy" and "attractor ruins", and also by W. Freeman (see, e.g.,[27]).

### IV. 'DEPOLARIZING' GABA ACTIONS - A POSSIBLE IMPLICATION

The non-classical, but frequently reported behaviour of GABA<sub>A</sub> receptor-mediated excitation in mature CNS has long been regarded as a puzzle [1], [20], [28], [18]. (See, also [15] and [29].)

In fact, a number of testimonies exist that GABA<sub>A</sub> have a depolarizing and excitatory action in *mature* CNS neurons. It is reported that when intensively activated, dendritic GABA<sub>A</sub> receptors excite rather than inhibit post-synaptic neurons, which are normally hyperpolarized by brief GABA<sub>A</sub> receptor activation. For example, Staley *et al.* ([28]) showed that while a single distal orthodromic stimuli induces hyperpolarizing GABA<sub>A</sub> receptor mediated PSPs (post-synaptic potentials), a train of 40 stimuli at 200 Hz yielded additional late depolarizing potential, and this depolarization is strong enough to modulate the voltage-dependent Mg<sup>2+</sup> block of synaptic NMDA receptor activations. Staley *et al.* [28] called this synapse-specific phenomenon as "activity-

dependent GABA<sub>A</sub> receptor-mediated depolarization", and emphasized that this should be distinguished with the rapid *activity-independent* GABA<sub>A</sub> depolarizing responses during developmental period. They also proposed a mechanism of such excitatory GABA actions (the anionic gradient shift model). We may abbreviate such activity-dependent excitatory GABA actions as "eGABA" in the following.

We give a hypothetical argument in [11] on the functional meaning of this eGABA actions from the standpoint of dynamic memory formation. We hypothesize that "eGABA" might act as a brain function to destabilize /stabilize memory traces to make them switch between Milnor and (semi-)classical attractors. Our arguments are based on, in contrary to an intuitive image, that GABAergic inhibitory systems may, under a certain condition, bring internal instability in the system. More precisely, inhibitions may destabilize attractors (as components of the memory structure in the brain). Here by the term "destabilization" we mean the "ruinization" or "milnor-ization" of the memory Evidence of such a structure under consideration. hypothetical argument may be found in e.g., Nara et al. [22], [23] and others though the situation is not exactly the same as our eGABA case. (See [11] for more discussions.)

A possible *scenario* is that a subsystem of GABAergic interneurons might play a *significant* or at least a *supplementary* role in constructing and re-organizing memory structures in collaboration with AMPA and NMDA channels. Then, GABA changes at consecutive stages its role to an *inhibitor*. This *switching* transforms the newly organized "attractor" into an "attractor ruin". The point here is that the same GABAergic interneuron system, which is one of the *concerned* or responsible components that brought the membrane potential to fire, changes its activity from excitatory to inhibitory. This fact may constitute the critical factor in this *ruin genesis*.

In view of the time course of eGABA activity (see, Kaila *et al.* [18]), it might also carry some cognitive role even at the stage of *perception*. Top down signals as "attention", "contextual information" etc. may have the possibility to extensively excite a certain subsystem of GABAergic systems. This is a highly hypothetical scenario, but might deserve further studies.

With a somewhat metaphoric expression, one might say that activation of eGABA actions temporarily changes the attractor *landscape* of the memory system by temporal organization of "attractors". When GABA turns its activity *back* to inhibitory, the structure which formed a "temporal attractor" breaks down to fragmental cell assemblies, each of which may exhibit a low activity chaotic state. However, as such temporarily linked fragmental cell assemblies cause correlated (or, synchronous) firings for some time, NMDA channels eventually open causing consolidation of a new memory structure. This newly-born structure turns to an attractor ruin in a background state.

It is an interesting supposition that top-down signals as category information or attention might play a critical role in such a dynamic memory formation via inhibitory "super networks".

#### V. CONCLUDING REMARKS

Some subgroups among a large variety of interneurons in the neocortex and hippocampus might form 'super networks' driven by top down signals as contexts or attentions (Buzsaki, 2001 [4]). This in turn may control and cooperatively entrain pyramidal networks which we presume possess multiple Milnor attractors as memory traces. Such a view on neocortical dynamics may give us a new perspective towards a unified framework on the cognitive basis of the neocortex and hippocampus.

We have discussed three controversial issues about the behaviour of interneurons which are presumably a part of such "super networks", and their possible consequences viewed from cognitive functions.

#### ACKNOWLEDGMENTS

The first author was supported by the Advanced and Innovative Research Program in Life Sciences and Grant-in-Aid no. 16500188 for Scientific Research (C), while the second author was supported by Grant-in-Aid no. 16650042 for Exploratory Research and also by Grant-in-Aid for formation of COE, all from Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government.

#### REFERENCES

- [1] B. E. Alger and R. Nicoll, Feed-forward dendritic inhibition in rat hippocampal pyramidal cell studied in vitro, **J. Physiology** (*Lond*) **328** (1982), pp. 105-123.
- [2] M. Blatow *et al.*, A Novel Network of Multipolar Bursting Interneurons generates Threta Frequency Oscillations in Neocortex, **Neuron 38** (2003), pp. 805-817.
- [3] G. Buzsaki, The hippocampo-neocortical dialogue, Cerebral Cortex 6 (1996), pp. 81-92.
- [4] G. Buzsaki, Hippocampal GABAergic Interneurons: A Physiological Perspective, **Neurochemical Res. 26** (2001), pp. 899-905.
- [5] B. Cauli *et al.*, Molecular and physiological diversity of cortical nonpyramidal cells, **J. Neurosci.** (1997), pp. 3894-3906.
- [6] J. A. Connor *et al.*, Neural repetitive firing: modifications of the Hodgkin-Huxley axon suggested by experimental results from crustacean axons, **Biophys. J. 18** (1977), pp. 81-102.
- [7] T. F. Freund, Interneuron Diversity Series: Rhythm and mood in perisomatic inhibition, **Trends in Neurosci. 26** (2003), pp. 489-495.

- [8] T. F. Freund and G. Buzsaki, Interneurons of the Hippocampus, **Hippocampus 6** (1996), pp. 347-470.
- [9] H. Fujii and I. Tsuda, Neocortical gap junction-coupled interneuron systems may induce chaotic behavior itinerant among quasi-attractors exhibiting transient synchrony, **Neurocomputing 58-60** (2004), pp. 151-157.
- [10] H. Fujii and I. Tsuda, Itinerant Dynamics of Class I\* Neurons Coupled by Gap Junctions, **Computational Neuroscience: Cortical Dynamics, Lecture Notes in Computer Science**, Vol. 3146, E'rdi, P.; Esposito, A.; Marinaro, M.; Scarpetta, S. (Eds.) 2004, pp. 140-160.
- [11] H. Fujii, K. Aihara and I. Tsuda, Functional Relevance of 'Excitatory' GABA Actions in Cortical Interneurons: a Dynamical Systems Approach, J. Integrative Neuroscience 3 (2004), pp. 183-205.
- [12] M. Galarreta and. S. Hestrin, A network of fast-spiking cells in the neocortex connected by electrical synapses, **Nature 402** (1999), pp. 72-75.
- [13] J. R. Gibson *et al.*, Two networks of electrically coupled inhibitory neurons in neocortex. **Nature 402** (1999), pp. 75-79.
- [14] C. Gray, A. K. Engel, P. Koenig and W. Singer, Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties, **Visual Neuroscience 8** (1992), pp. 337-347.
- [15] A. T. Gulledge and G. J. Stuart, Excitatory Actions of GABA in the Cortex, Neuron 37 (2003), pp. 299-309.
- [16] J. L. Hindmarsh and R. M. Rose, A model of neuronal bursting using three coupled first order differential equations, **Proceedings of Royal Society of London, B221** (1984), pp. 87-102.
- [17] A. L. Hodgkin, The local electric changes associated with repetitive action in a non-medullated axon, **J. Physiology 107** (1948), p. 165-181.
- [18] K. Kaila *et al.*, Long Lasting GABA-Mediated Depolarization Evoked by High-Frequency Stimulation in Pyramidal Neurons of Rat Hippocampal Slice Is Attributable to a Network-Driven, Bicarbonate-Dependent K<sup>+</sup> Transient, **J. Neurosci. 17** (1997), pp. 7662-7672.
- [19] H. Markram *et al.*, Interneurons of the Neocortical Inhibitory System, **Nature Rev. Neurosci. 5** (2004), pp. 793-807.
- [20] H. Michelson and R. K. S. Wong, Excitatory synaptic responses mediated by  $GABA_A$  receptors in the hippocampus, **Science 253** (1991), pp. 1420-1423.
- [21] J. Milnor, On the concept of attractor, Communications in Mathematical Physics 99 (1985), pp. 177-195.
- [22] S. Nara and P. Davis, Chaotic wandering and search in a cycle-memory neural network, **Progress of Theoretical Physics 88** (1992), pp. 845-855.
- [23] S. Nara *et al.*, Chaotic memory dynamics in a recurrent neural networks with cycle memories embedded by pseudo-inverse method. **Int. J. of Bifurcation and Chaos 5** (1995), pp. 1205-1212.
- [24] J. Rinzel, Excitation dynamics: insights from simplified membrane models, Fed. Proc. 44 (1985), pp. 2944-2946.
- [25] M. A. Rogawski, The A-current: how ubiquitous a feature of excitable cells is it? **Trends in Neurosci. 8** (1985), pp. 214-219.
- [26] R. M. Rose and J. L. Hindmarsh, The assembly of ionic currents in a thalamic neuron I. The three-dimensional model, **Proc. R. Soc. Lond. B 237** (1989), pp. 267-288.

- [27] C. Skarda and W. J. Freeman, How brains make chaos in order to make sense of the world, **Behavioral and Brain Sciences 10** (1987), pp. 161-195.
- [28] K. J. Staley, B. L. Soldo and W. R. Proctor, Ionic Mechanisms of Neuronal Excitation by Inhibitory  $GABA_A$  Receptors, Science 269 (1995), pp. 977–981.
- [29] V. Stein and R. A. Nicoll, GABA generates excitement, Neuron 37 (2003), pp. 375-378.
- [30] I. Tsuda, Chaotic itinerancy as a dynamical basis of Hermeneutics of brain and mind, **World Futures 32** (1991), pp. 167-185.
- [31] I. Tsuda, Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems, **Behavioral and Brain Sciences 24** (2001), pp. 793-847.
- [32] I. Tsuda, H. Fujii, S. Tadokoro, T. Yasuoka and Y. Yamaguti, Chaotic Itinerancy as a Mechanism of Irregular Changes between Synchronization and Desynchronization in a Neural Network, **J. Integrative Neuroscience 3** (2004), pp. 159-182.
- [33] I. Tsuda and H. Fujii, A Complex Systems Approach to an Interpretation of Dynamic Brain Activity I: Chaotic itinerancy can provide a mathematical basis for information processing in cortical transitory and nonstationary dynamics, Computational Neuroscience: Cortical Dynamics, Lecture Notes in Computer Science, Vol. 3146, E'rdi, P.; Esposito, A.; Marinaro, M.; Scarpetta, S. (Eds.) 2004, pp. 109-128.
- [34] X.-J. Wang. Synaptic Basis of Cortical Persistent Activity: the Importance of NMDA Receptors to Working memory, **J. Neurosci 19** (1999), pp. 9587-9603.